Supplementary data

Molecular cloning of Aed a 3

The salivary gland λgt11 cDNA library of adult mosquito *Ae. aegypti* (provided by Dr. A. James, University of California, Irvine) was screened on *E. coli* Y1090 at a density of approximately 20,000 PFU per 150 mm plates. After 3.5 hr incubation at 42°C, nitrocellulose filters (0.45 am) (Habersham, Chicago, IL, USA), impregnated with 10 mm IPTG, were applied to the agarose plates and incubated at 37°C for 3.5 hr. Each filter was rinsed in TBST (50 mM Tris, pH7.9, 150 mM NaCl, 0.05% Tween-20) prior to overnight incubation at 4°C with 20% fetal bovine serum (FBS) in TBST to block free protein binding sites. The filters were then incubated for 1.5 hr at room temperature with 1:1,000 of mouse anti-saliva serum probe in TBST implemented with 1% BSA (TBST-BSA), washed three times over 30 min in TBST, and incubated for 1 hr at room temperature with a 1:3,000 dilution of alkaline phosphatase-conjugated goat anti-mouse IgG (Jackson ImmunoResearch Lab. Inc., West Grave, PA, USA) in TBST-BSA. After washing again, the filters were stained with BCIP (5-bromo-4-chloro-3-indoylphosphate p-toluidine salt) and NBT (p-nitro blue tetrazolium chloride) substrate (Bio-Rad, Hercules, CA, USA). Positive plaques were removed from the agarose plate and subjected to subsequent rounds of purification on 85 mm plates until all of the plaques reacted with the mouse anti-saliva serum.

5’ RACE PCR was designed to clone the 5' terminal fragment from the cDNA library using a sense primer, 5' GACTCCTGGAGCCCG 3', (λgt11 forward primer) (Clontech, Palo Aito, U.S.A) and an antisense primer, 3' ATATCTGTCCACCAACG 5', which is complementary to a sequence of the 3' terminal fragment. PCR was performed in a 100-µl sample containing 0.5 µg of library DNA, 10 µl of 10 × buffer, 2 µl of 25 mM dNTP's, 2 µl of 100 ng/µl each primer, and 2 U of Taq polymerase supplied by the PCR kit (Boehringer Mannheim Canada, Quebec). The reaction was subjected to 25 cycles of amplification consisting of 1 min at 94°C, 1 min at 55°C and 1 min at 72°C, with a final 72°C extension for 7 min. The PCR products were cloned into the PCR™ TA vector (Invitrogen, San Diego, CA, USA). Mini-preparation of plasmid DNA was employed to isolate the PCR products, which were subject to restriction enzyme digestion and sequencing analysis. Two sets of PCR primer pairs were designed to combine two overlapped cDNA clones into a full-length Aed a cDNA clone. Set 1 was for the 5’ clone: 5' primer

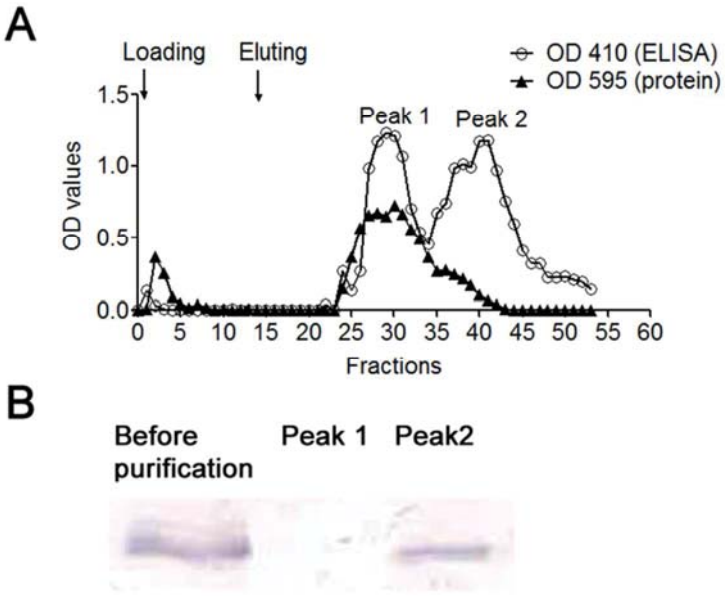
(5'ATGGATCCTCCGAAAATGAAACCCTTG3') and 3' primer (5'CGAACCACCTGTCTATA3'). Set 2 was for the 3' clone: 5' primer (5'TGAAAAAGAGGAAGGAG3') and 3' primer (5'CCGGTACCCGTTACAGACAGAATGAGG3'). The PCR products were purified by GENECLEAN II Kit. After being denatured, these two fragments were annealed by their overlapped region, and filled in by DNA polymerase. This full-length template was amplified by PCR using the above 5' primer and the 3' primer. The PCR reaction was performed in a 50 µl volume: 5 µl of 10X PCR buffer, 1 µl of 10 mM dNTP, 30 pmol of each primer, 2.5 U of Pwo polymerase. The reaction was subjected to 20 cycles of amplification for 5' and 3' fragment isolation, or 30 cycles for full-length amplification. The full-length clone was confirmed by sequencing.

Expression and purification of rAed a 3

The full-length Aed a 3 cDNA was cloned into baculovirus expression vector pVL1393 by BamHI and PstI sites. Recombinant baculoviruses carrying the pVL1393/Aed a 3 cDNA were generated by cotransfection of S. frugiperda (Sf9) insect cells with the transfer vector pVL1392/Aed a 3 together with linearized baculovirus AcMNPV DNA using the cationic liposome method. Pure recombinant viral stocks were prepared as described previously.[1, 2] To increase the yield of the expressed protein, High-Five cells cultured in EX-cell™ 405 media (JRH Biosciences) were used to replace the Sf9 cells. After infection with the recombinant virus and assessing the time course of expression, cultures were harvested and then centrifuged twice.

In the purification procedure, concentrated rAed a 3 supernatant was dialyzed against Tris buffer (10 mM Tris. Buffer, pH 7.5) and then loaded onto a DEAE-Sephacel column equilibrated with Tris buffer. After washing the column with Tris buffer, the bound proteins were eluted with a linear gradient NaCl from 0 to 0.6 M in Tris buffer. Fractions were monitored for protein concentrations using OD 595 nm and assayed for rAed a 3 content levels using ELISA. In the ELISA, microplates coated with fractions (1:5) were incubated with mouse anti-rAed a 3 serum (previously prepared with an unpurified rAed a 3 fraction in our laboratory) followed by incubation with alkaline phosphatase-conjugated goat anti-mouse IgG. Since the ELISA results showed two peaks after the gradient elution, Western blot was performed to see which peak contains rAed a 3. The results showed that peak 2, not peak 1, contained rAed a 3 (**Supplementary Figure 1**).

Fractions in peak 2 were pooled and concentrated with a final protein concentration of 0.2 mg/ml for use of skin tests and serum rAed a 3-specific antibody assay.



Supplementary figure 1. Purification of rAed a 3.

A. Chromatography on DEAE-Sephacel. Concentrated rAed a 3 supernatant was loaded onto a DEAE-Sephacel column. The bound proteins were eluted with a linear gradient NaCl from 0 to 0.6 M in Tris buffer. Protein concentrations were monitored using OD_{595} nm and rAed a 3 content levels were assayed using ELISA. Since the ELISA results showed two peaks after the gradient elution, Western blot was performed to see which peak contains rAed a 3.

B. Westernblot analysis. Fractions of peak 1 and peak 2 were pooled and concentrated separately. Proteins separated on 12% SDS-PAGE were transferred to nitrocellulose membrane and immunoblotted with mouse anti-mosquito saliva serum.

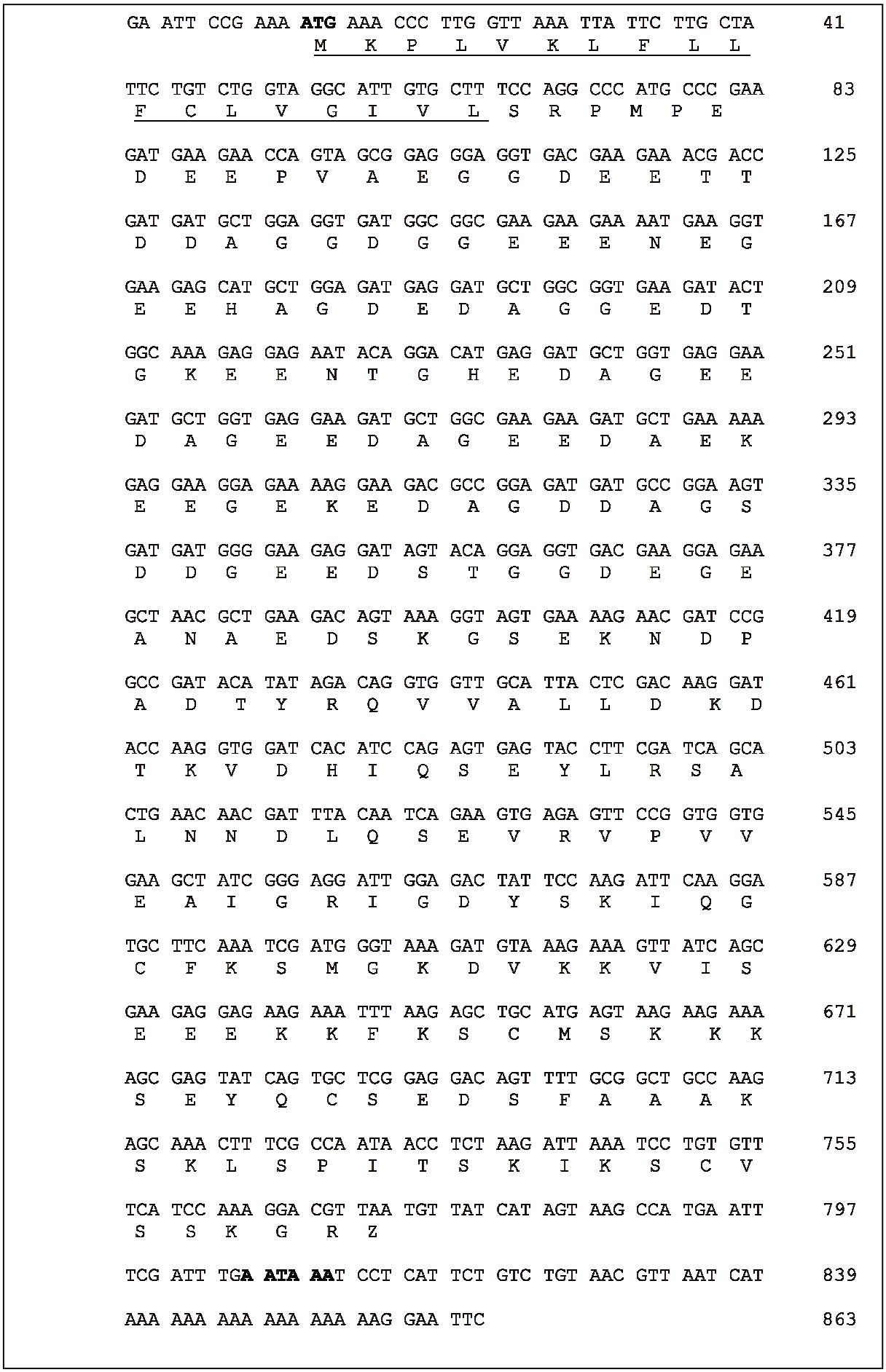
Preparation of fusion protein-selected antibodies

In order to prepare antibodies specific to the cloned gene's protein, nitrocellulose membrane absorbed with the protein produced by the cloned gene were used to purify specific antibodies from mouse anti-mosquito saliva serum and also from mosquito-allergic human serum using the method previously reported.[3] Briefly, the selected clone from λgt11 cDNA library screening was infected into *E. coli* Y1089 as a lysogen expressing the cloned gene's protein. These lysogens were plated onto 85 mm agar plates at a high density so that plaques were nearly confluent at the end of incubation. Nitrocellulose filters, impregnated with 10 mM IPTG, were overlaid onto the lawn of *E. coli* 1090 in the agar plate and incubated for 5 hr at 37°C for gene expression induction. Each filter was washed and blocked for 2 hr at room temperature in 3% BSA in TBST and then incubated overnight at 4°C with 10 ml of mouse antiserum of mosquito-allergic human serum (diluted 1:100 in TBST-BSA). Antibodies were eluted by incubating the filter for 10 min with 5 ml of glycine buffer (0.1 M, pH 2.8). The antibodies were aspirated and neutralised to pH 7 by adding 1 M Tris buffer (pH 8.0). These clone-specific antibodies were used to pinpoint its original protein contained in the saliva of *Ae. aegypti* using immunoblotting.

cDNA cloning and sequence analysis

Screening of about 120,000 plaques from the whole cDNA library of adult mosquito *Ae. aegypti* with mouse anti-saliva serum identified 39 positive plaques. All the 39 clones were purified by 3 rounds of plaque selection and were grouped into A, B and C according to the color strength of their reactions to mouse anti-serum on the filters. Four clones of each group were chosen to prepare fusion proteins by infecting *E. coli* Y 1089 strain with the cloned phages. Twelve samples of the fusion proteins were immunoblotted by mouse anti-saliva serum. A protein band with molecular weight range from 125 to 175 kDa was found in each sample (data not shown). As the β-galactosidase part of the fusion protein is 114 kDa, the cDNA coding part ranged from 10 to 60 kDa. Three clones with different sizes of cDNA inserts were subcloned into pBluescript II SK vector and then checked by sequencing both the 5' and 3' ends of the inserts. One clone (λAA22) having a complete 3' terminus with poly A tail was selected for further analysis. This fragment was determined to contain the major coding region of the 30 kDa protein as the cloned cDNA (0.73 kb) has an open reading frame for 217 amino acids, the molecular weight of which is estimated to be 26 kDa. In the 3' terminus, the features consistent with known eukaryotic genes, a consensus polyadenylation sequence, AATAAA, and polyadenosines were observed. Since the cDNA insert was expressed as a fusion protein with ß-galactosidase, the cDNA coding protein must be in the same frame with ß-galactosidase protein. An open reading frame for 217 amino acids was observed in the frame with the ß-galactosidase protein. This clone missed the 5' terminal sequence because no initiation codon was observed in the 5' terminus of the clone.

PCR was designed to clone the 5' terminal fragment of the full-length cDNA. PCR products were cloned into a TA vector and 3 clones were obtained and sequenced. The sequences of the 3 clones were found to be identical and overlapped λAA22. In the search for an initiation codon in the same open reading frame as λAA22 cDNA, an ATG was found, which contained an adenosine at the crucial -3 position of the Kozak consensus sequence, A/GXXXATG, for initiation of translation by eukaryotic ribosomes.[4] The sequence flanking the putative translational start sites, GAAAATG, is outstandingly identical to the consensus sequence, C/AAAA/CATG, for initiation of *Drosophila*, an insect gene.[5] The full-length cDNA is 0.85 kb, coding for a protein of 254 amino acid residues or approximately 30 kDa. The primary nucleotide sequence of the Aed a 3 cDNA and its deduced amino acid sequence have been indicated in **Supplementary data figure 2**. The sequence data have been deposited in the GenBank databases under accession No. AF001927 in 1997.



```
GA ATT CCG AAA ATG AAA CCC TTG GTT AAA TTA TTC TTG CTA      41
               M   K   P   L   V   K   L   F   L   L

TTC TGT CTG GTA GGC ATT GTG CTT TCC AGG CCC ATG CCC GAA     83
 F   C   L   V   G   I   V   L   S   R   P   M   P   E

GAT GAA GAA CCA GTA GCG GAG GGA GGT GAC GAA GAA ACG ACC     125
 D   E   E   P   V   A   E   G   G   D   E   E   T   T

GAT GAT GCT GGA GGT GAT GGC GGC GAA GAA GAA AAT GAA GGT     167
 D   D   A   G   G   D   G   G   E   E   E   N   E   G

GAA GAG CAT GCT GGA GAT GAG GAT GCT GGC GGT GAA GAT ACT     209
 E   E   H   A   G   D   E   D   A   G   G   E   D   T

GGC AAA GAG GAG AAT ACA GGA CAT GAG GAT GCT GGT GAG GAA     251
 G   K   E   E   N   T   G   H   E   D   A   G   E   E

GAT GCT GGT GAG GAA GAT GCT GGC GAA GAA GAT GCT GAA AAA     293
 D   A   G   E   E   D   A   G   E   E   D   A   E   K

GAG GAA GGA GAA AAG GAA GAC GCC GGA GAT GAT GCC GGA AGT     335
 E   E   G   E   K   E   D   A   G   D   D   A   G   S

GAT GAT GGG GAA GAG GAT AGT ACA GGA GGT GAC GAA GGA GAA     377
 D   D   G   E   E   D   S   T   G   G   D   E   G   E

GCT AAC GCT GAA GAC AGT AAA GGT AGT GAA AAG AAC GAT CCG     419
 A   N   A   E   D   S   K   G   S   E   K   N   D   P

GCC GAT ACA TAT AGA CAG GTG GTT GCA TTA CTC GAC AAG GAT     461
 A   D   T   Y   R   Q   V   V   A   L   L   D   K   D

ACC AAG GTG GAT CAC ATC CAG AGT GAG TAC CTT CGA TCA GCA     503
 T   K   V   D   H   I   Q   S   E   Y   L   R   S   A

CTG AAC AAC GAT TTA CAA TCA GAA GTG AGA GTT CCG GTG GTG     545
 L   N   N   D   L   Q   S   E   V   R   V   P   V   V

GAA GCT ATC GGG AGG ATT GGA GAC TAT TCC AAG ATT CAA GGA     587
 E   A   I   G   R   I   G   D   Y   S   K   I   Q   G

TGC TTC AAA TCG ATG GGT AAA GAT GTA AAG AAA GTT ATC AGC     629
 C   F   K   S   M   G   K   D   V   K   K   V   I   S

GAA GAG GAG AAG AAA TTT AAG AGC TGC ATG AGT AAG AAG AAA     671
 E   E   E   K   K   F   K   S   C   M   S   K   K   K

AGC GAG TAT CAG TGC TCG GAG GAC AGT TTT GCG GCT GCC AAG     713
 S   E   Y   Q   C   S   E   D   S   F   A   A   A   K

AGC AAA CTT TCG CCA ATA ACC TCT AAG ATT AAA TCC TGT GTT     755
 S   K   L   S   P   I   T   S   K   I   K   S   C   V

TCA TCC AAA GGA CGT TAA TGT TAT CAT AGT AAG CCA TGA ATT     797
 S   S   K   G   R   Z

TCG ATT TGA ATA AAT CCT CAT TCT GTC TGT AAC GTT AAT CAT     839

AAA AAA AAA AAA AAA AAG GAA TTC                            863
```

Supplementary figure 2. Primary nucleotide sequence of Aed a 3 cDNA and its deduced amino acid sequence

From the searches of the DNA and protein databases to determine the identity of the cloned cDNA, it is apparent that this putative protein represents a novel protein. Although the BLASTN results based on the nucleotide sequence indicate a high degree of similarity to a number of known sequences, these similarities most likely result from a number of repetitive codons in the sequence (data not shown). This interpretation is confirmed by the BLASTP search based on the conceptual translation product of the cDNA, which does not indicate any similarity to a known protein. A more sensitive FASTA search showed some extent of local sequence similarity of this putative protein to a number of known proteins. These include herpes virus 3 α trans-inducing factor, plasmodium falciparum glutamic acid-rich protein, etc. This novel protein showed a 43.69% identity in 119 amino acid overlap to the herpes virus Saimiri hypothetical gene 48 protein.[6] Whether these proteins have similar biological functions is unclear. The protein is rich in glutamic acids (16.5% of amino acid residues), and has a hydrophobic amino terminal region characteristic of a secretory signal peptide. Searching Aed a 3 against motif database Pfam did not give any hit, however the more sensitive Prosite database gave 22 hits including a Glu-rich region and several Protein kinase C phosphorylation sites, Casein kinase II phosphorylation sites and N-myristoylation sites. The biological functions of this protein were later investigated as described in the discussion section of the paper.

References

1. Peng Z, Xu W, James AA, Lam H, Sun D, Cheng L, et al. Expression, purification, characterization, and clinical relevance of rAed a 1 - a 68 kDa recombinant mosquito *Aedes aegypti* salivary allergen. Inter Immunol 2001; 13:1445-52.
2. Xu W, Simons FER, Peng Z. Expression and rapid purification of an Aedes aegypti salivary allergen by a baculovirus system. Int Arch Allergy Immunol 1998; 115:245-51.
3. Caraballo L, Avjioglu A, Marrugo J, Puerta L, Marsh D. Cloning and expression of complementary DNA coding for an allergen with common antibody-binding specificities with three allergens of the house dust mite Blomia tropicalis. J Allergy Clin Immunol 1996; 98:573-9.
4. Kozak M. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res 1987; 15:8125-48.

5. Cavener DR. Comparison of the consensus sequence flanking translational start sites in Drosophila and vertebrates. Nucleic Acids Res 1987; 15:1353-61.

6. Albrecht JC, Nicholas J, Biller D, Cameron KR, Biesinger B, Newman C, et al. Primary structure of the herpesvirus saimiri genome. J Virol 1992; 66:5047-58.